**Electric Vehicle Data Analysis Project Plan**

**Project Overview**

* **Objective**: Analyze the performance and specifications of various electric vehicle (EV) models.
* **Goals**:
  + Find trends in EV performance (e.g., top speed vs. battery capacity)
  + Analyze market segments (e.g., SUV vs sedan)
  + Study relationships (e.g., range vs. efficiency)
  + Provide insights for manufacturers or customers

**Environment Setup**

python

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

**Step 1: Load the Data**

# adjust file path as needed

df = pd.read\_csv("electric\_vehicles\_spec\_2025.csv.csv")

# preview

df.head()

**Step 2: Understand Data Types & Shape**

# basic info

print(df.info())

# shape of the dataset

print(f"Rows: {df.shape[0]}, Columns: {df.shape[1]}")

**Step 3: Check Missing Values**

# count missing values per column

missing = df.isnull().sum()

missing[missing > 0]

**Step 4: Descriptive Statistics**

# numeric columns

print(df.describe())

# categorical columns

print(df.describe(include='object'))

**Step 5: Data Cleaning**

Here you might do things like:

* fill missing numeric data with mean/median
* fill missing categorical data with mode
* remove duplicates

# remove duplicates if any

df = df.drop\_duplicates()

# example: fill missing torque\_nm with median

df['torque\_nm'] = df['torque\_nm'].fillna(df['torque\_nm'].median())

# example: fill missing car\_body\_type with mode

df['car\_body\_type'] = df['car\_body\_type'].fillna(df['car\_body\_type'].mode()[0])

**Step 6: Exploratory Data Analysis (EDA)**

**a) Correlation heatmap**

plt.figure(figsize=(12,8))

sns.heatmap(df.corr(numeric\_only=True), annot=True, cmap='coolwarm')

plt.title("Correlation Heatmap")

plt.show()

**b) Distribution plots**

# battery capacity

sns.histplot(df['battery\_capacity\_kWh'], kde=True)

plt.title("Battery Capacity Distribution")

plt.show()

**c) Top Brands by Range**

brand\_range = df.groupby('brand')['range\_km'].mean().sort\_values(ascending=False).head(10)

brand\_range.plot(kind='barh', color='green')

plt.xlabel("Average Range (km)")

plt.title("Top Brands by Average Range")

plt.show()

**d) Segment Distribution**

df['segment'].value\_counts().plot(kind='pie', autopct='%1.0f%%', figsize=(8,8))

plt.title("Segment Share")

plt.ylabel("")

plt.show()

**Step 7: Advanced Feature Relationships**

**Top speed vs. range**

sns.scatterplot(x='top\_speed\_kmh', y='range\_km', data=df, hue='car\_body\_type')

plt.title("Top Speed vs. Range by Car Body Type")

plt.show()

**Battery capacity vs. efficiency**

sns.scatterplot(x='battery\_capacity\_kWh', y='efficiency\_wh\_per\_km', data=df, hue='segment')

plt.title("Battery Capacity vs. Efficiency")

plt.show()

**Step 8: Business Insights**

After the graphs, you could summarize insights like:

* *“SUVs have the highest average towing capacity.”*
* *“Models with higher battery capacity generally achieve better range, but their efficiency varies.”*
* *“Fast charging power is strongly correlated with top speed.”*

**Step 9: Reporting**

Summarize in a slide deck or Jupyter markdown:

* **Project Goal**: analyzing EV specs
* **Key Findings** (from charts & correlations)
* **Recommendations** (e.g., where to focus for improvement)
* **Limitations** (missing data, outdated data, etc.)

**EV Data Analysis in Tableau — Step-by-Step**

**Step 1: Load your CSV in Tableau**

1. Open **Tableau Desktop** (or Tableau Public, free)
2. Click **Connect to Data → Text File**
3. Select your electric\_vehicles\_spec\_2025.csv.csv file
4. Tableau will preview your data
5. Click **Sheet 1** to begin your analysis

**Step 2: Clean and inspect your data**

* Use Tableau’s **Data Source** tab to:
  + Check column data types (change from string to number if needed)
  + Fix any nulls or unexpected formats
  + Rename columns for clarity if you like

**Step 3: Build Visualizations**

Here are some ideas for visuals:

**1. Average range by segment**

* Drag segment to **Columns**
* Drag range\_km to **Rows**
* Change aggregation to **AVG**
* Add color encoding for better visuals

**2. Battery capacity distribution**

* Use a histogram on battery\_capacity\_kWh

**3. Top brands by range**

* brand on **Rows**
* average range\_km on **Columns**
* sort descending

**4. Scatterplot of top speed vs range**

* drag top\_speed\_kmh to X-axis
* drag range\_km to Y-axis
* put car\_body\_type on **Color**

**5. Segment distribution**

* segment on **Rows**
* Number of Records on **Columns**
* convert to Pie chart

**6. Box plot of acceleration by car body type**

* acceleration\_0\_100\_s on **Columns**
* car\_body\_type on **Rows**
* change chart type to box-and-whisker

**Step 4: Add Calculated Fields (if needed)**

For example, to calculate **cargo volume in m³**:

* Right-click in data pane → *Create Calculated Field*
  + Name: Cargo Volume m3
  + Formula: [cargo\_volume\_l] / 1000

**Step 5: Build a Dashboard**

* Add multiple sheets to a **dashboard**
* Use filters (e.g., select a car brand and see details update)
* Add legends, titles, and interactivity

**Step 6: Publish / Share**

* If using Tableau Public → **Publish** to Tableau Public profile
* If using Tableau Desktop → share the .twbx packaged workbook